

Euro Working Group on Transportation Annual Meeting 2025 - EWGT2025

Integrating demand data with train delay models: A socio-economic evaluation for maintenance planning

Abderrahman Ait-Ali^{a,b}, Zohreh Ranjbar^c, Martin Joborn^{b,c}

^aCommunications and Transport Systems, Linköping University, Norrköping, Sweden

^bTransport Economics, VTI Swedish National Road and Transport Research Institute, Stockholm, Sweden

^cMobility and Systems, RISE Research Institutes of Sweden, Västerås, Sweden

Abstract

Railway punctuality remains a critical measure of service quality and operational efficiency. Traditional performance metrics, such as on-time performance and delay increments, inform about punctuality goals and guide maintenance planning, but they often overlook the passenger experience due to limited access to disaggregated demand data. This study integrates forecasted ridership data with delay evaluation models to assess passenger delays and their socio-economic impacts. By combining passenger-centric delay contributions with the Swedish framework for socio-economic evaluations, we enable a more informed prioritisation of maintenance interventions. A case study on the Southern Main Line in Sweden illustrates the methodology's potential to improve maintenance planning, highlighting its relevance for achieving data-driven improvements in train service reliability.

© 2026 The Authors. Published by ELSEVIER B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the Euro Working Group on Transportation Annual Meeting 2025 - EWGT2025.

Keywords: Railway punctuality; passenger delay; socio-economic cost; ridership estimation; maintenance planning.

1. Introduction

Railway service reliability and punctuality are among the most important performance indicators for both passenger satisfaction and infrastructure management. Punctuality, often monitored using vehicle-centric measures, remains the most widely used indicator in operational practice (Goverde, 2005). In Sweden, these indicators are not only used for performance monitoring and contractual penalties, but also for planning maintenance activities on important railway infrastructure assets. For instance, train disruption hours, i.e., accumulated incremental delays, serve as one basis for prioritisation when maintenance plans are prepared; the infrastructure manager prioritises interventions based on locations and asset types contributing most to delays, see a review by Pettersson (2020) for Swedish railways.

However, existing approaches for punctuality and performance evaluation primarily focus on train-centric metrics, overlooking passenger-level impacts, in particular on congested time intervals and track sections. This gap can lead to

discrepancies between perceived and actual service quality, as ridership patterns, delay propagation, and heterogeneities in service usage are not fully accounted for.

To address these limitations, this study incorporates ridership estimate data into delay models to better evaluate passenger delays and their socio-economic costs. The focus is therefore on disturbed passenger trains. The long-term effects of delays on passenger demand, as shown by Nelldal et al. (2022), are not included in the evaluations.

The approach in this work builds on the concept of delay contributions, introduced by Joborn and Ranjbar (2022), and adapts it to a passenger-centric and socio-economic framework. A case study on the Southern Main Line in 2023 demonstrates the value of this methodology for delay assessment, maintenance planning and prioritisation.

The paper is structured as follows. Section 2 reviews related literature on punctuality metrics, socio-economic evaluations, and maintenance planning. Section 3 presents the proposed methodology, including the integration of forecasted ridership data and delay models. Section 4 describes the case study, results and key findings, followed by conclusions and directions for future research in Section 5.

2. Literature review

2.1. Punctuality metrics

In most European (passenger) railway systems, punctuality is typically evaluated through vehicle-centric metrics, e.g., using vehicles' arrival times at terminal stations (van Oort, 2014). For instance, the Swedish infrastructure manager (Trafikverket) uses RT+X as the standard punctuality metric for monitoring national railways, indicating the share of trains arriving within X minutes of scheduled arrival time at specific checkpoints. Complementary to this, delay increments greater than three minutes between specific measurement points are used in performance regimes (Odolinski, 2019). Delays exceeding five minutes (or cancellations) can trigger financial penalties under these regimes (Nilsson, 2016). Such metrics are therefore widely adopted for both national performance reporting and contractual agreements between stakeholders (Kristoffersson and Pyddoke, 2019).

These metrics are well-suited for system-level reporting and accountability, but are limited in their ability to capture passenger experience, especially at intermediate (large) stations. Moreover, they treat all (empty or full) trains equally, regardless of passenger load, peak/off-peak hours, and/or trip purpose, potentially skewing reliability monitoring, infrastructure investment and operational priorities. This limitation has motivated the development of complementary metrics, see a recent review by Ait-Ali (2024) exploring alternative passenger-centric indicators. For instance, Joborn and Ranjbar (2022) introduced delay contribution, which focuses on the causal analysis of delays with respect to their impact on final punctuality outcomes. Another example is the use of more passenger-centric metrics for monitoring punctuality and service reliability, which are becoming increasingly used thanks to the availability of new automatic data sources for counting passengers, e.g., Automatic Fare Collection (AFC) and Automatic Passenger Counting (APC) systems (Hendren et al., 2015).

2.2. Socio-economic evaluation

Another important complementary approach for following up punctuality is the socio-economic evaluation of delays. Socio-economic evaluations are central in transport planning and project appraisal in several countries, e.g., Sweden, where these evaluations are guided by national ASEK guidelines (Broberg and Wettergren, 2024). These guidelines provide monetary value for travel time and delay costs per passenger, differentiated by trip purpose and train type. Typically, higher delay costs are assigned to business trips (e.g., 19 SEK per pax-min) compared to private or commuting trips (e.g., 4-5 SEK per pax-min), according to the latest guidelines from ASEK 8.0 (Broberg and Wettergren, 2024). These valuations are used in, among others, cost-benefit analyses (CBA) for strategic investment decisions and (to a lesser extent) in tactical operation planning, e.g., capacity allocation (Warg et al., 2019, Ait-Ali et al., 2020). However, these are rarely integrated into delay evaluation models, punctuality monitoring or maintenance planning.

Few previous studies have explored the use of socio-economic evaluations to assess train delays. Nelldal et al. (2022) looked at delay impacts on demand and modal choice, highlighting that persistent delays may reduce long-term ridership. Most of the existing studies focused on road delays (Bivina et al., 2016). However, few studies have

addressed the application of socio-economic assessment at the level of individual disruptions or infrastructure failures. Integrating such valuation into operational models could enable more accurate and cost-effective prioritisation of interventions, e.g., (preventive) maintenance planning.

2.3. Maintenance planning

Rail infrastructure maintenance is generally divided into corrective and preventive maintenance. The latter consists of periodic scheduled-based activities and condition-based interventions that are based on the results from periodic monitoring of the infrastructure assets (Sund and Canaki, 2025).

Prioritisation between different interventions, as part of preventive maintenance planning, is traditionally guided by performance metrics and expert judgment to minimise system disruptions and maintain safety while considering available budget limits (Lidén, 2014). In Sweden, train disruption hours—aggregated from delay increments—are used to monitor performance and plan maintenance budgets annually (Pettersson, 2020). These disruption hours serve as an input for identifying problematic locations and asset types (e.g., switches, rail, signalling systems), which are then prioritised for preventive interventions based on their cumulative impact on train punctuality and budget restrictions.

However, current planning frameworks typically lack integration of demand-side information. As such, disruptions affecting a small number of high-load trains may not be prioritised, despite potentially large socio-economic costs. This gap between performance metrics and passenger experience creates inefficiencies in the allocation of (limited) maintenance resources. Bridging this gap requires methodologies that account for both train-level disruptions and their impact on passengers, particularly in critical segments and assets of the railway network.

2.4. Summary

Previous related studies have explored railway delay and punctuality modelling, passenger demand estimation, or railway maintenance planning, but few attempts have been made to bridge these topics. Delay modelling has evolved from using simple (vehicle-centric) metrics to more passenger-centric indicators and advanced concepts, e.g., delay contributions and critical disturbances. Socio-economic evaluations, such as CBA, are common in policy analyses but rarely operationalised for day-to-day prioritisation in infrastructure maintenance. Passenger demand, when considered, is often aggregated and disconnected from operational maintenance planning.

While the reviewed literature has contributed to improved monitoring of the performance of railway operations with/without demand data, none of the existing reviewed studies, to our knowledge, offer a framework that integrates the assessment of individual disruptions using socio-economic principles within an operational context such as maintenance planning. This study addresses this gap by combining train delay modelling, demand forecasts, and socio-economic evaluation to assess passenger delays and support maintenance prioritisation.

3. Methodology

The methodology comprises three core components, which are illustrated schematically in Fig. 1, namely train load estimation, delay assessment and socio-economic cost evaluation.

3.1. Train load estimation

Most European railway markets, including Sweden, have been vertically separated, i.e., into infrastructure management and train operations (Ait-Ali and Eliasson, 2021). Infrastructure managers have, therefore, limited access to data such as passenger demand, even if APC and AFC data collection systems are increasingly deployed, at least for publicly operated local systems. Hence, the need for the train load estimation, which is, in this work, based on demand prognosis. If passenger data is available, e.g., via APC and/or AFC, this train-load estimation step can be skipped.

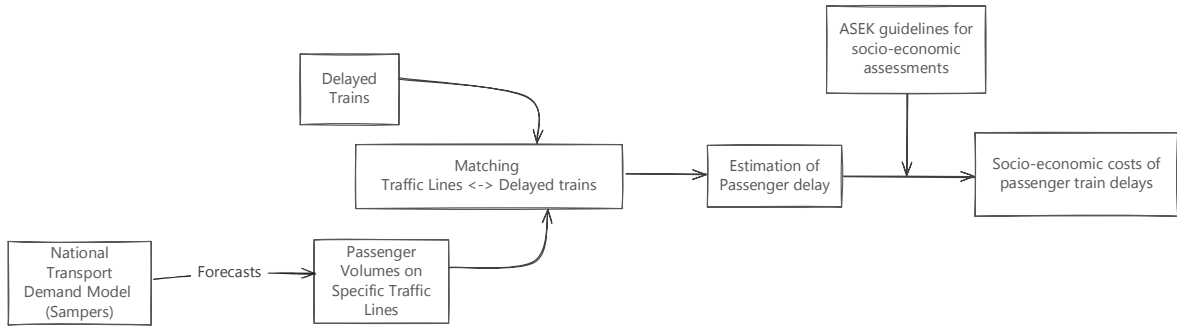


Fig. 1. Overview of the methodology illustrating some of the main steps in this study.

To overcome this, passenger volumes on different traffic lines are obtained from the national transport demand forecast model (*Sampers*), which estimates the daily average passenger volumes on specific traffic lines (Trafikverket, 2020). To obtain an estimation of affected passengers, delayed trains are matched to the most likely traffic line, based on their stopping patterns.

Each delayed train $t \in T$ is matched to its most likely traffic line $l \in L$ based on a similar stopping pattern. The train-line matching function is based on a similarity score which considers, among others, the similarity of intermediate stops, first/last stops, and the number of stops (in case of a tie).

The similarity score function $S(t, l)$ is defined in Equation (1), where α_i are selected weights (e.g., emphasising first/last and intermediate stops $\alpha_1 = \alpha_3 > \alpha_2$), and sim_x is a measure of overlapping stops, e.g., Jaccard index, focusing on first stop (for $X = first_stop$), on intermediate stops (for $X = int_stops$), on last stop (for $X = last_stop$).

$$S(t, l) = \alpha_1 \cdot sim_{first_stop}(t, l) + \alpha_2 \cdot sim_{int_stops}(t, l) + \alpha_3 \cdot sim_{last_stop}(t, l) \quad (1)$$

Average daily passenger volumes from the matched traffic line P_l are used to estimate the average affected passengers P_t alighting from the delayed trains, i.e., using the line with the highest similarity score, see Equation (2). More detailed estimates of the number of affected passengers P_t are available differentiated by train type and trip purpose, $P_{t,i}$ where $i \in I$, e.g., business, leisure/private or work/commuting for national or regional trains.

$$P_t = P_{argmax_l S(t,l)} \quad (2)$$

Note that the focus in this study is on the passengers *alighting* from the delayed trains. However, similar analyses can be performed with a focus on *boarding* passengers, *onboard* passengers, or a *weighted combination* of all, depending on the evaluation context, i.e., whether the focus is on departure/arrival delay, and/or transfers. Also, note that P_t are static average volumes per train, i.e., no dynamic variation over time, e.g., peak/off-peak.

3.2. Delay assessment

Passenger delays are assessed by integrating ridership data into delay models. Passenger-level delays are calculated by multiplying the estimated ridership by the delay duration of each delayed train. While train-hour delays are commonly used, as mentioned earlier, this study explores the use of passenger-centric metrics such as passenger-hour delays, offering a more nuanced analysis.

Let $T(d)$ be the set of disturbed trains because of disturbance $d \in D$, Equation (3) computes the total passenger delay PD_d associated with disturbance d . It is also possible to calculate the passenger delay per train type and trip purpose $PD_{d,i}$ by considering the corresponding number of affected passengers $P_{t,i}$ where $i \in I$. To integrate the delay contribution concept (Joborn and Ranjbar, 2022) and reflect passenger-centric impacts, the parameter Δ_t represents the delay contribution affecting train $t \in T(d)$ because of disturbance d . This can facilitate, e.g., the identification of infrastructure-related disturbances d with high values of passenger delays PD_d .

$$PD_d = \sum_{t \in T(d)} P_t \Delta_t \quad (3)$$

3.3. Socio-economic cost evaluation

In economic appraisals, travel times are differentiated for different train types and trip purposes, e.g., business, private or commuting. The former is typically evaluated as the highest. Travel delay times are therefore assigned different cost values and are differentiated per trip purpose.

Passenger delays are converted into monetary costs using values of travel time from, e.g., the Swedish ASEK 8.0 framework (Broberg and Wettergren, 2024). The delay cost parameters v_i , which is 3.5 times the value of travel time, assigns different values per passenger depending on the train type and trip purpose $i \in I$. The total socioeconomic costs are calculated as in Equation (4).

$$C_d = \sum_{i \in I} v_i \cdot PD_{d,i} \quad (4)$$

This monetary valuation facilitates the comparison of societal costs across different disturbances, enabling infrastructure managers to prioritise high-impact maintenance interventions based on socio-economic efficiency, e.g., causing disturbances d with very large socioeconomic costs C_d .

4. Case study and results

To illustrate the application of the proposed methodology, we conduct a case study on the Southern Main Line (*Södra Stambanan*) between two major Swedish cities (*Stockholm* and *Malmö*). A train trip over the entire line is around 616 km and takes, according to the timetable, on average 4 h and 27 min (Joborn and Ranjbar, 2022). This line is one of the most critical corridors in the Swedish railway network, connecting the capital and the southern region of the country. It serves a combination of long-distance, regional, and freight traffic, and is used by commuter trains with relatively high passenger volumes, especially during peak hours. These reasons, and the availability of relevant delay and traffic data, motivate the choice of the selected corridor.

4.1. Data

The analyses are based on multiple data sources, including traffic and delay data, as well as passenger demand estimates. The analysis focuses on the selected corridor during the timetable year 2023.

Delay and disturbances data were retrieved from different Trafikverket's monitoring systems such as *Lupp*, *Bessy* and *Ofelia*. The final dataset includes, among others, reported failures in the corridor and their attributed causes, as well as registered disturbances, e.g., the number of disturbed trains and their delays. Each disturbance is associated with one or more delayed trains and is classified by type, such as infrastructure-related, external, vehicle-related, etc.

In total, the dataset covers more than 9 thousand (disturbed) passenger trains, of which more than 5 thousand are delayed, i.e., more than five minutes to their final station. These delays are due to different types of failures, of which more than two thousand (around 36% of all failures) are infrastructure-related failures. Passenger demand data were obtained from *Sampers*, which provides estimates of average daily ridership volumes by traffic line. The demand data are disaggregated by train type (regional and national) and trip purpose (commuting, business, and private/others). This disaggregation enables differentiated valuation in later steps. Around 22,500 passengers alight daily at some of the major stations on the studied corridor. Table 1 presents some descriptive statistics from the datasets used.

Table 1. Descriptive statistics from disturbance, delay and demand data.

Category	Description	Value
Disturbances	Disturbed trains in the dataset	9 300
	Total infrastructure-related failures (% of the total)	2 220 (36%)
Delay	Delayed trains, i.e., more than 5 min to the final station	5 200
	Average delay duration in minutes	15
Demand	Total number of daily alighting passengers	22 490
	Average (and standard deviation) per station	460 (1 040)

Socio-economic valuation of delays is based on cost parameters from the national Swedish guidelines (Broberg and Wettergren, 2024). The guidelines provide the standard delay cost values for different trip purposes and types, which are used to convert passenger delays into monetary costs. By combining the estimated number of passengers per train, delay durations, and trip type and purpose distributions, it is possible to compute the aggregated total socio-economic cost for the different individual disturbances.

4.2. Results

The analysis of delays and their socio-economic costs reveals variations across track sections and/or asset types. First, we compare two measures, namely incremental delay, which reflects only primary delays exceeding three minutes, and passenger delay, which captures the cumulative impact of delay propagation and dissipation along the journey, weighted by passenger alighting volumes. The results reveal some discrepancies between the two measures, underscoring the importance of passenger-centric delay assessment.

Fig. 2(a) shows that high-impact delays are due to infrastructure-related failures that are heavily concentrated in specific track sections, e.g., the section between *Åby södra* and *Norrköping central* is responsible for around 25% of total infrastructure-related delays. The results highlight some discrepancies between the two metrics. For instance, several track sections, e.g., between *Katrineholm* and *Åby södra*, display relatively low incremental delays but much higher passenger delays. This reflects the phenomenon where delays either propagate or dissipate as trains progress, as well as differences in passenger ridership across disturbed trains. These dynamics are entirely missed by incremental delay analyses alone.

Fig. 2(b) shows the distribution of the metrics across different asset types. Clearer discrepancies emerge, e.g., while signalling systems are responsible for the highest share of incremental delays, the overhead lines (*Kontaktledning*) generate the largest share (nearly 30%) of total passenger delays. This highlights how certain asset failures, even not the most frequent/long-lasting, can have significant passenger impacts due to where/when they occur. Other major contributors to passenger delays include switches and crossings (*Spårväxel*) and signalling systems. These results show the importance of considering passenger delay rather than purely technical delay when evaluating maintenance priorities, as asset categories that seem less critical from a train delay perspective may carry much higher societal costs when passenger impacts are accounted for.

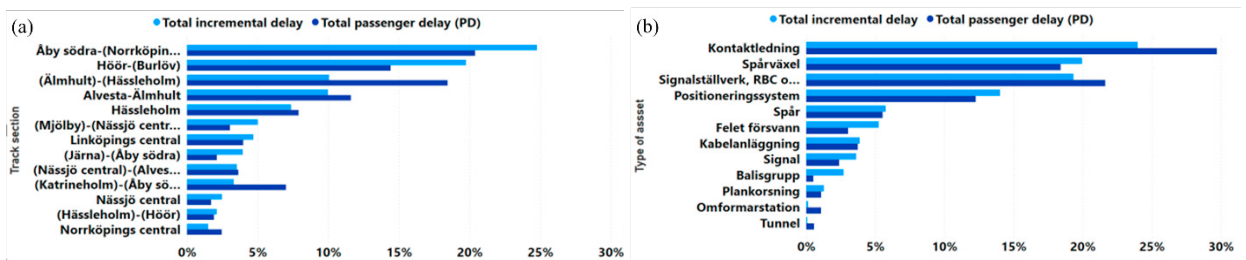


Fig. 2. Distribution (in %) of total incremental delay and total passenger delay (PD) across (a) track sections, and (b) asset types.

Socio-economic cost results are presented in Fig. 3 as a heatmap/matrix of the monetary impact of delays on passengers, based on delay contribution data at every station along the route, combined with passenger alighting forecasts. The matrix reveals critical hotspots where delays translate into substantial societal costs. Notably, overhead lines (*Kontaktledning*) issues along the *Älmhult–Hässleholm*, and the *Åby södra–Norrköping–Linköping–Mjölby* section. Such analyses are suitable, e.g., for identifying infrastructure components (what) and/or track sections (where) that have incurred high socio-economic costs.

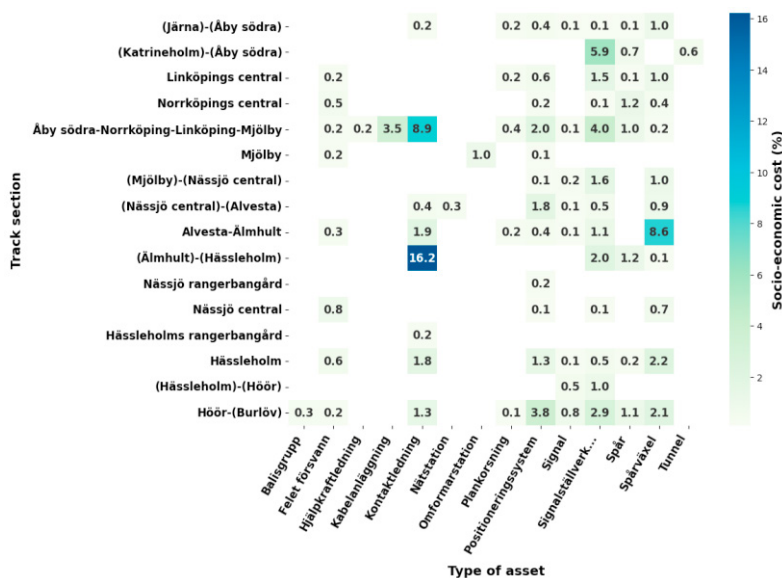


Fig. 3. Socio-economic costs of delays (in %) disaggregated by track sections and asset types. Darker shades represent higher socio-economic cost shares. Key cost drivers include coupling systems, cabling installations, and signalling equipment, with particularly high impacts observed on the *Ålhmult–Hässleholm*, *Åby södra–Norrköping–Linköping–Mjölby* and *Alvesta–Ålhmult* sections.

4.3. Discussions

The results reveal disparities between traditional incremental delay measurements and passenger-centric delay assessments (or the associated socio-economic costs). Comparison across different track sections shows that specific sections stand out as major causes of both large incremental and passenger delays. However, several sections with only moderate incremental delays, such as those with high alighting passenger volumes, generate disproportionately high passenger delays. At the asset level, the differences between the two perspectives are clearer. Some assets account for the largest shares of incremental delays. Yet, when passenger delays are considered, other assets emerge as the largest causes. This discrepancy underscores the importance of considering passenger movement patterns when assessing the impact of technical failures. Failures that occur in locations and/or times where many passengers are alighting, even if technically minor, can have major impacts in socio-economic terms.

The study of the associated socio-economic costs of delays, across both track sections and asset types, is useful to highlight hotspots where infrastructure-related failures lead to significant passenger costs. These patterns/hotspots would be invisible or distorted if relying solely on incremental delay metrics. It is, however, important to note that passenger alighting data were not directly observed but estimated using a national transport model, as described in the methodology. While this introduces some uncertainty, it represents a substantial improvement over analyses that assume uniform or unknown passenger distributions, as well as delay assessment at the *train’s final* station.

The results presented in this study are specific to one corridor and rely on forecasted demand data. While this limits generalizability, the framework itself is adaptable. The model can be applied to other corridors using either forecasted or observed demand (if available) and can be extended to allow for more dynamic travel behaviour, such as passengers rerouting or rescheduling trips in response to frequent delays.

5. Conclusions

This work has demonstrated how integrating ridership forecasts can enhance the evaluation of infrastructure-related disturbances in large railway networks. It has also introduced passenger-centric delay assessment and associated socio-economic costs as alternative metrics for operational and maintenance decision-support. By focusing on passenger

experience and societal costs rather than solely on operational metrics, IMs can achieve more effective maintenance prioritisation and contribute to a more resilient and reliable railway.

Potential extensions include the use of improved demand data (e.g., time-dependent ridership, real-time or observed occupancy data), dynamic demand modelling under disruptions, and application of the framework to include freight services.

While the proposed methodology offers a new approach to integrate ridership forecasts into delay valuation models, several limitations should be acknowledged. First, the estimation of train-specific passenger volumes relies on forecasted demand data and a line-matching algorithm based on stopping patterns. This may introduce uncertainties in the accuracy of assigned ridership, e.g., special trains or trains with deviations from typical patterns. This could be resolved by using observed demand data from automatic collection systems, e.g., automatic passenger counting (APC), automatic fare collection (AFC) systems or mobile data. Using such data would enhance the accuracy of demand estimates. Second, the passenger behaviour is assumed to remain constant regardless of service disruptions. Significant or recurring delays may influence travel patterns, e.g., modal shifts. Incorporating these dynamic behaviours would further refine the model. Third, while this study focused on a single critical corridor, testing across larger networks, as well as over multiple years, would be valuable for more insights into the scalability and robustness of the approach. Finally, future work could explore how the resulting costs vary under different assumptions for passenger demand, delay propagation.

Acknowledgements

This work was supported by the Swedish Transport Administration (Trafikverket) through the KAJT research program. The authors are grateful to two anonymous reviewers for their useful comments on an earlier version of this paper.

References

- Ait-Ali, A. 2024. Exploring Measures to Monitor Passenger-Centric Punctuality in Public Transport. *K2 Working Papers 2024:9*.
- Ait-Ali, A. & Eliasson, J. 2021. European Railway Deregulation: An overview of market organization and capacity allocation. *Transportmetrica A: Transport Science*, 1-27.
- Ait-Ali, A., Warg, J. & Eliasson, J. 2020. Pricing commercial train path requests based on societal costs. *Transportation Research Part A: Policy and Practice*, 132, 452-464.
- Bivina, G. R., Landge, V. & Kumar, V. S. S. 2016. Socio Economic Valuation of Traffic Delays. *Transportation Research Procedia*, 17, 513-520.
- Broberg, T. & Wettergren, H. B. 2024. Analysmetod och samhällsekonomiska kalkylvärden för transportsektorn: ASEK 8.0.
- Goverde, R. M. 2005. *Punctuality of railway operations and timetable stability analysis*, Netherlands TRAIL Research School Delft.
- Hendren, P., Antos, J., Carney, Y. & Harcum, R. 2015. Transit Travel Time Reliability: Shifting the Focus from Vehicles to Customers. *Transportation Research Record*, 2535, 35-44.
- Joborn, M. & Ranjbar, Z. 2022. Understanding causes of unpunctual trains: Delay contribution and critical disturbances. *Journal of Rail Transport Planning & Management*, 23, 100339.
- Kristoffersson, I. & Pyddoke, R. A traveller perspective on railway punctuality: Passenger loads and punctuality for regional trains in Sweden. RailNorrköping 2019. 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköping, Sweden, June 17th–20th, 2019, Linköping University Electronic Press, 565-578.
- Lidén, T. 2014. *Survey of railway maintenance activities from a planning perspective and literature review concerning the use of mathematical algorithms for solving such planning and scheduling problems*.
- Nelldal, B.-L., Näsman, P. & Song, H.-S. 2022. Vidareutveckling av effektsamband för förseningarnas påverkan på efterfrågan av tågresor. KTH Järnvägsgrupp.
- Nilsson, J.-E. 2016. *Kvalitetsavgifter: problem och tänkbara lösningar*, Statens väg-och transportforskningsinstitut.
- Odolinski, K. 2019. Contract design and performance of railway maintenance: Effects of incentive intensity and performance incentive schemes. *Economics of Transportation*, 18, 50-59.
- Pettersson, T. 2020. En punktligare tågtrafik : sammanställning av Trafikverkets åtgärder 2017–2019. *Trafikverkets publikationer*. Borlänge: Trafikverket.
- Sund, J. & Canaki, A. 2025. Trafikverkets underhållsplan : för åren 2025–2028. *Trafikverkets publikationer*. Borlänge: Trafikverket.
- Trafikverket 2020. Sampers och trafikprognoser : en kort introduktion. *Trafikverkets publikationer*.
- van Oort, N. 2014. Incorporating service reliability in public transport design and performance requirements: International survey results and recommendations. *Research in Transportation Economics*, 48, 92-100.
- Warg, J., Ait-Ali, A. & Eliasson, J. 2019. Assessment of Commuter Train Timetables Including Transfers. *Transportation Research Procedia*, 37, 11-18.